

# EOSC Multi-Annual Roadmap 2023-2024

## **Acknowledgement**

The EOSC Association Board wishes to thank its Task Forces for their immense contributions under a tight timeframe to ensure the initial draft of the MAR was prepared in time for community consultation. We also wish to thank all those who took time to read the document and respond to the consultation survey. Many constructive comments were received and have enabled us to produce a much more robust roadmap for 2023-2024. Finally we wish to thank the EOSC Association secretariat including Ute Gunsenheimer (Secretary General) and René Buch (CTO) for assistance in shaping the MAR, and particularly Giulia Lodi (Communications Officer) for ensuring the text was placed into the appropriate template for release to Members and Observers before the General Assembly.

Brussels, May 2022



## Table of Contents

<b>1. Introduction</b> .....	<b>4</b>
<b>2. Background context</b> .....	<b>5</b>
<b>2.1 Results of the MAR open consultation</b> .....	<b>5</b>
<b>2.2 Minimum Viable EOSC (MVE)</b> .....	<b>8</b>
2.2.1 Scope and timing of the MVE.....	10
<b>2.3 H2020 and Horizon Europe EOSC projects</b> .....	<b>12</b>
<b>3. Roadmap 2023-2024</b> .....	<b>14</b>
<b>3.1 SRIA objectives</b> .....	<b>14</b>
<b>3.2 Levels of implementation</b> .....	<b>16</b>
<b>3.3 Summary of priorities</b> .....	<b>16</b>
<b>3.4 Objective 1 - Open Science as the new normal</b> .....	<b>18</b>
3.4.1 European level priorities.....	18
3.4.2 National level priorities.....	19
3.4.3 Institutional level priorities.....	19
3.4.4 Expected outcomes.....	20
3.4.5 Clarification text.....	20
<b>3.5 Objective 2 - Adoption of standards and tools</b> .....	<b>22</b>
3.5.1 European level priorities.....	22
3.5.2 National level priorities.....	23
3.5.3 Institutional level priorities.....	23
3.5.4 Expected outcomes.....	24
3.5.5 Clarification text.....	24
<b>3.6 Objective 3 - Establish a sustainable and federated infrastructure</b> .....	<b>26</b>
3.6.1 European level priorities.....	26
3.6.2 National level priorities.....	27
3.6.3 Institutional level priorities.....	27
3.6.4 Expected outcomes.....	27
3.6.5 Clarification text.....	28
<b>4. Conclusion</b> .....	<b>29</b>



## List of Figures

Figure 1.	The three phases of EOSC implementation	4
Figure 2.	Categories of stakeholders responding to the consultation	5
Figure 3.	Stakeholder response capacity	5
Figure 4.	Distribution of agreement and appropriateness for objectives and outcomes	6
Figure 5.	Percentage of respondents who fundamentally disagree with aspects of the MAR	7
Figure 6.	The Minimum Viable EOSC	9
Figure 7.	Scope and timing of the MVE	11
Figure 8.	EOSC projects active since 2021	13

## 1. Introduction

This Strategic Research and Innovation Agenda (SRIA) sets clear goals to develop the European Open Science Cloud (EOSC) and build a research environment that promotes Open Science and increases trust in and reproducibility of research. The overall impact is a pan-European research landscape that offers significantly improved discovery, access, interoperability, and exploitation of research outputs for researchers and for research and innovation stakeholders.

The Horizon Europe (HE) funding periods are mirrored in the SRIA implementation stages (see Figure 1). Each stage defines a high-level objective for the period, aligned with the vision of enabling a trusted, virtual, federated environment in Europe to store, share and reuse research outputs across borders and scientific disciplines. For each stage, priority areas are defined, their prioritisation and scheduling reflecting community feedback from the open consultation process.

### **Stage 1 (2021–2022): Development towards added value from a functional federation of infrastructures**

Enabling the European Open Science Cloud operations (the EOSC-Core) to provide necessary core functions of the Minimum Viable EOSC (MVE) that allow federation of existing and future infrastructures, according to associated Rules of Participation and governance that provision growth and expansion in the following stages.

### **Stage 2 (2023–2024): Expansion to production that generates added value**

Expanding and building the core data infrastructure to support the full lifecycle of scientific research in key thematic areas. During this period, activities will build on pilots / demonstrators and work towards linking EOSC beyond the research communities to the wider public sector and the private sector.

### **Stage 3 (2025–2027 and beyond): Expansion to develop impact from Open Science**

Deployment of federated research infrastructures with functionality that allows multiple European research communities to deliver impactful Open Science. In addition to European infrastructures, the national research infrastructures delivered from the Member States and Associated Countries in particular will help in this expansion phase.

#### Figure 1. The three phases of EOSC implementation

This section proposes a work plan for the 2023-2024 period, taking into account the activities planned for the first stage and those forthcoming under the Horizon Europe INFRAEOSC Work Programme. This Multi-Annual Roadmap (MAR) 2023-2024 has been drafted by the EOSC Association Board of Directors in collaboration with the EOSC Association's Task Forces. The content has also been shaped by a community consultation held in March 2022, which provided inputs on priorities. The roadmap provides a framework of priorities with derived activities and indicators for the second stage of activities. These focus on embedding EOSC by increasing the number of Research Infrastructures (RIs) that are federated and growing the range of research outputs and services that are available. Activities also focus on developing skill sets, offering support networks and increasing the adoption of standards, so Open Science and FAIR practices are more widely adopted. The EOSC user base should be widened through collaboration with other Partnerships, data spaces and public sector groups. The

EOSC is seen as one of the Data Spaces that overlaps with all other data spaces for their Research and Innovation activities. Ultimately, by the 2025-2027 period of development, the various thematic data spaces in development should connect via EOSC.

## 2. Background context

### 2.1 Results of the MAR open consultation

The open community consultation on the draft MAR 2023-2024 resulted in a large number of valuable comments that have been addressed and incorporated into this final MAR 2023-2024. The EOSC Association is immensely grateful to the community for the time taken to review the draft document and provide such meaningful, considered comments.

The first draft of the MAR for 2023-2024 was released on Tuesday, 1st March 2022 and the consultation period ran for three weeks, closing on Monday, 21st March 2022. Over that period, 45 complete responses were received, resulting in 534 individual comments. Most of the responses came from Research Performing Organisations (RPOs) and service providers (see Figure 2). Most also responded on behalf of the stated organisation rather than in an individual capacity (see Figure 3).

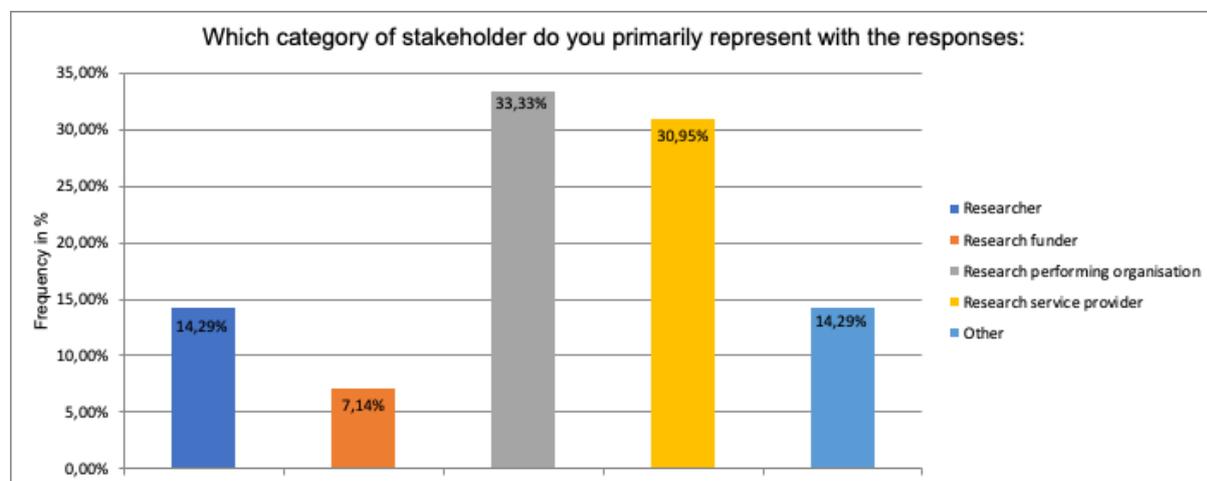


Figure 2. Categories of stakeholders responding to the consultation

In which capacity are you responding?

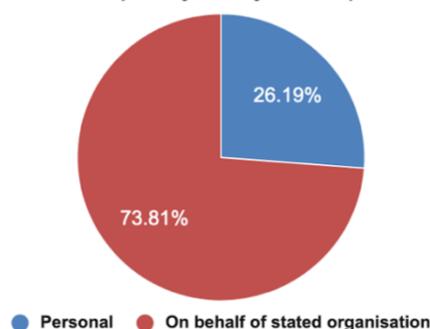
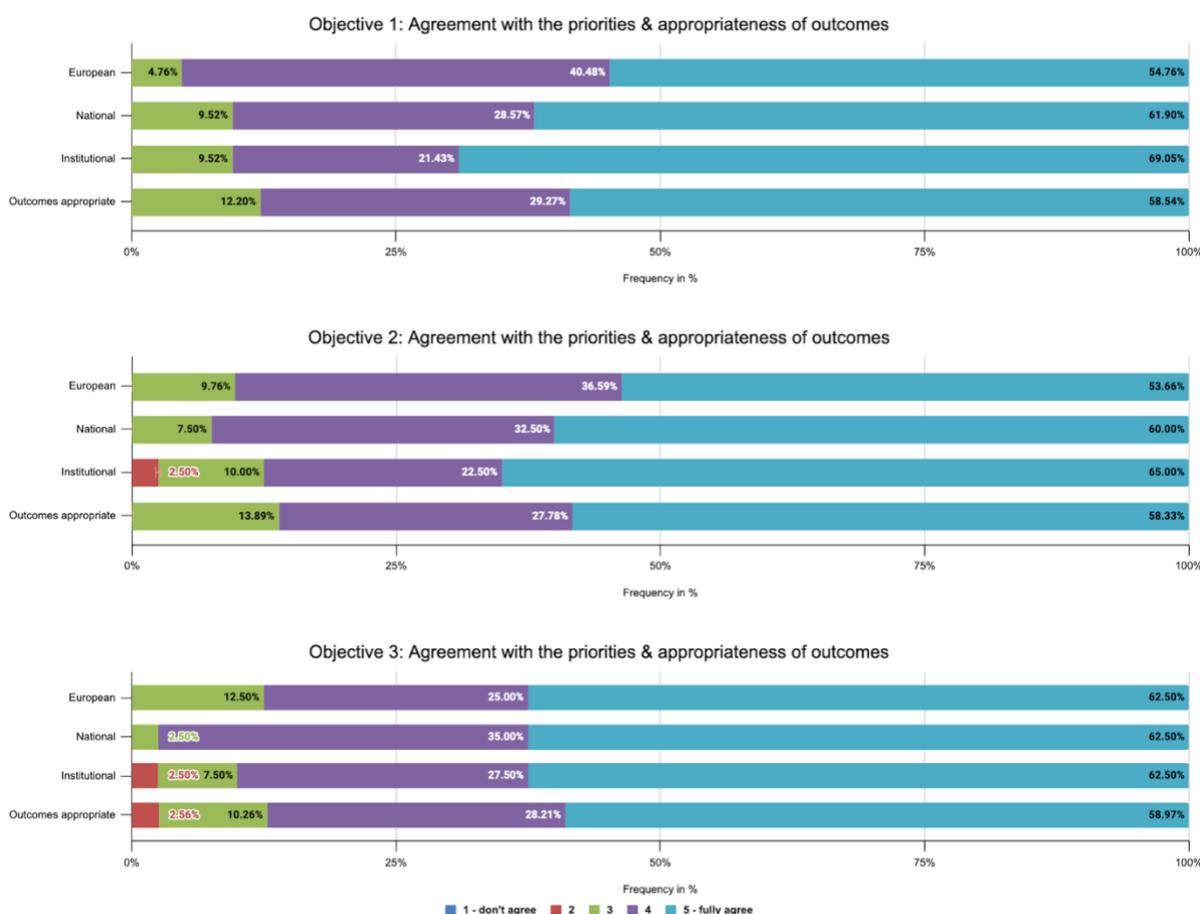


Figure 3. Stakeholder response capacity

The survey included a number of structured questions to gauge the level of agreement with the priorities defined, as well as allowing specific comments and suggestions for changes. Overall, there was a high-level of acceptance with an average of 60% of the community being in full agreement with the priorities and outcomes for each objective. Nobody was in full disagreement for any of the priorities and outcomes (see Figure 4). The comments typically alluded to small textual amendments or changes of emphasis and prioritisation which have been addressed in this final version.



**Figure 4. Distribution of agreement and appropriateness for objectives and outcomes**

When asked if there was any aspect in the MAR which they fundamentally disagreed with and felt should be removed or changed, only 16% of respondents said yes (see Figure 5). Five comments were left on this point and related to several important aspects which we have addressed. There were a number of concerns about how EOSC-related activities will be funded and a fear that if the resources are fixed then money will be taken away from research activity. A suggestion was made to pursue a greater degree of national co-funding as is done by other Joint Undertakings such as EuroHPC and this is precisely the intention within EOSC. One commenter noted an over-emphasis on the EOSC Portal and a lack of cross-border funding models to ensure the “marketplace” is viable. This is the top priority as the technical implementation cannot function without effective resourcing models in place.

Are there fundamental aspects you disagree with, which you feel should be removed / amended?

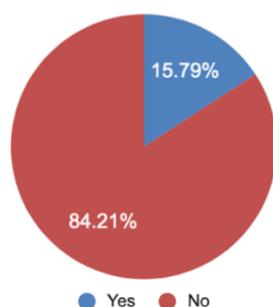


Figure 5. Percentage of respondents who fundamentally disagree with aspects of the MAR

Others were concerned about reinventing wheels and not using existing community standards and Persistent Identifier (PID) mechanisms. As this is an unfounded concern, text changes have been made to emphasise the commitment to build on existing research community practices and models from Research Infrastructures. Another commenter felt that the MAR was biased towards data and did not address what is needed from the eInfrastructures to host, compute, and publish data. While we understand this concern, it is a challenging aspect to address since such services are competitive and fall with the EOSC-Exchange rather than EOSC-Core. There is a limit to the stimulation and support we can demand for such services.

Another area where the influence of the EOSC Association is constrained is in relation to funded projects. One commenter noted that the EOSC Association Task Forces are mainly open but largely supported by volunteer effort and unfunded, whilst the EOSC Future project is well funded and closed. We agree that all Open Science infrastructure projects should practise what they preach, regularly sharing information early and ensuring all software source code and other outputs are well-documented, conform to accepted standards, and are released under an open licence. This is critical to ensure public investments can continue to be built on by the community and developed into mature, sustainable services, irrespective of who first created them or is funded to continue work. The commenter also noted a disconnect in the relationship between the Association and the EOSC Future project. Several meetings have been held between the EOSC Association Board and EOSC Future representatives to promote collaboration towards a common goal and progress is slowly being made. However, under Horizon 2020, projects have no obligation to collaborate or share information with the EOSC Association - such practice is dependent on goodwill from both sides, with support from the European Commission (EC).

The final comment related to the structure of the MAR, noting that it seemed like a collection of bottom-up inputs rather than a coherent, structured set of priorities. This is also valid and is a consequence of the process by which we solicited inputs from the Task Forces. A significant amount of work has been done to refine the MAR, making the text clearer and sharper as requested, removing or merging overlapping priorities and defining a few key areas of work which need to be emphasised in the 2023-2024 period. This emphasis is important as there were also general comments noting that the roadmap was a little over-ambitious and unrealistic to achieve by the end of 2024. We have tempered some outcomes but on the whole prefer stretch targets to ensure we achieve as much as possible within the time frame.

A number of common themes emerged from the comments. These were classified, grouped and assigned to the most relevant Board member to address in consultation with the Task Force co-chairs, as needed. Most have been processed, though some, such as the value proposition and some terminological issues, are better placed in the wider SRIA text so will be fully addressed when that is updated in 2023. Changes should be evident in the revised text and the Association Board hopes that the community feels its voice has been heard and reflected. These common themes were as follows:

- Better articulate the value proposition for researchers to engage in EOSC and for RIs, institutions and other providers to federate their data and services
- Clarify terminology such as EOSC-Core, EOSC-Exchange, onboarding, federation and Stakeholder Forum
- Extend the priorities related to multilingualism
- Place more emphasis on national investment and the role of Member States
- Clarify how EOSC will be funded
- Prioritise business models and resourcing otherwise EOSC will not work
- Explain the role of data stewards and RSEs, who does the job and how we fund them
- Address concerns about reinventing wheels in the implementation of Authentication and Authorisation Infrastructure (AAI), PIDs and EOSC Interoperability Framework

The full consultation data will be released to the community so it is clear what comments were received and how these have been considered and addressed by the Board. At the time of writing we still need to identify which comments we have permission to make public and retract any for which this has not been granted.

## 2.2 Minimum Viable EOSC (MVE)

A functioning Minimum Viable EOSC (MVE) can only emerge from an active and coordinated community effort. The relevant stakeholders for co-creating EOSC are wide and varied, including H2020 and Horizon Europe projects, national and institutional data initiatives, research communities, funders, standards bodies and many others. The EOSC Association is responsible for stakeholder engagement and representing the voice of the community, however the relevant stakeholders go beyond its Members and Observers. The so-called “EOSC Stakeholder Forum” as referenced in the EOSC Executive Board Governance diagram released in 2020, is not a predefined group, body or a single event like the EOSC Symposium. Many fora, events and methods are used to engage and consult with stakeholders, each time reaching a different set of individuals. For this reason, we have chosen not to use the term “Stakeholder Forum” anymore to avoid confusion in the community. All stakeholders are relevant and can engage by whichever methods suit them. Coordination of engagement activities and stakeholders is essential and will be led by the Association.

The MVE shall deliver on the core SRIA objectives and provide foundational support for Open Science in Europe. Based on the definition made by the previous EOSC Architecture Working Group and the diagrams emerging from the EOSC Executive Board, the MVE includes four components (see Figure 6):

- **EOSC-Core**, which comprises the enabling services required to operate the EOSC

- **Federated data**, whereby metadata on research outputs is harvested into a cross-search to enable greater discovery and reuse of data residing in multiple institutional, domain-specific and national repositories across Europe
- **EOSC Interoperability Framework** which provides the guidelines, specifications, standards and APIs for the composition of EOSC services and resources
- **EOSC-Exchange**, which is composed of common and thematic services exploiting FAIR data and encouraging its reuse

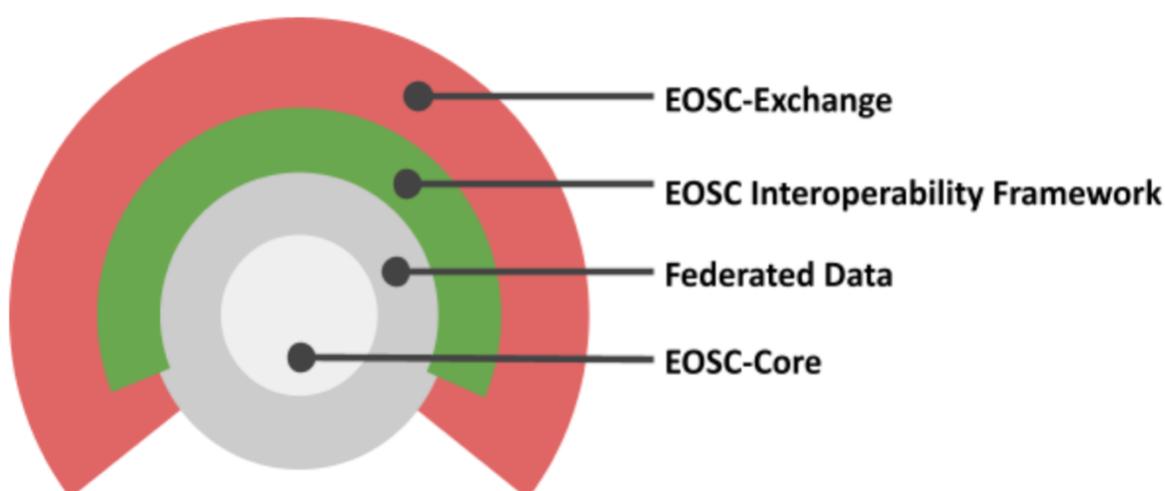


Figure 6. The Minimum Viable EOSC

Within EOSC, we are pursuing a system of systems approach, federating existing services and data. This means that it is not a single, monolithic infrastructure, rather an overarching platform which connects data and services, which reside in multiple locations across Europe by means of a series of accepted standards, crosswalks, and open APIs defined in the EOSC Interoperability Framework. There are a number of common services in the EOSC-Core, such as AAI and monitoring, to which service providers can connect their systems. This ensures users have seamless login and access via the EOSC. The federated model means that no services or users will be forced to change their existing practices. The AAI mechanism implemented by research infrastructures, for example, will still remain operational, but by ensuring they are compatible with EOSC standards such as the AARC Blueprint, these login mechanisms will be compatible with EOSC and enable access to a broader range of data and services offered by other infrastructures using other login mechanisms. Supporting the many languages in use across Europe via the multilingualism activities will also facilitate the federation of data and services. For these reasons, the EOSC Interoperability Framework is the critical connecting component. Federation is very time-consuming and complicated work so sufficient funds need to be provided to encourage and enable this transition.

The number, complexity, and interdependency of existing projects, initiatives, and involved stakeholders illustrates the ongoing challenge of building EOSC and raising awareness to a point where Open Science becomes the new normal. Horizon Europe will build on the results of existing projects both for core functionality and for thematic activities centred around federated data infrastructures.

As the voice of the community, the EOSC Association plays a pivotal role here. It oversees the implementation of the strategic objectives defined in this SRJA and will take over results from

the EOSC projects, where assessments performed by the community and the EOSC Association Board recommend they are fit for purpose. The EOSC Association will stimulate the coordination of activities and stakeholders to progress the MVE towards a functional and performant federated data infrastructure. For this to be effective, existing projects and forthcoming Horizon Europe projects will have to work on convergence, together with the EOSC Association. For the forthcoming projects this should be written into their grant agreements. In any case the help of the European Commission will be key to achieve the right reporting structure.

As noted, the development of an operational EOSC platform uses a federated approach to connect existing infrastructures. Present work brings together the infrastructures with all five research clusters (ENVRI-FAIR, EOSC Life, PANOSC, SSHOC, and ESCAPE) to ensure the integration of data and resources from their communities. A set of processes to onboard services into EOSC Portal has been defined in terms of usage, access policies, working with resource providers, privacy policies, and other technical and administrative procedures. In the 2023-2024 period, these procedures should be aligned with the Rules of Participation (RoP) which were defined by the EOSC Executive Board and are being continued by the EOSC Association Task Force. The Rules of Participation state the conditions, quality, information, and service levels of the contributors to EOSC so provide the critical boundary conditions for engagement.

### *2.2.1 Scope and timing of the MVE*

The staged approach to the development of EOSC described in the SRIA is presented in Figure 7. The approach covers the key required areas and development periods to enable EOSC engagement over a broad community, extending beyond researchers to the private sector. The period of development described in the SRIA covers two periods of EC funding: Horizon 2020 and Horizon Europe. It is expected that EOSC will be delivered through the activities of a large number of engaged stakeholders. Stage 1 describes the activities to be initiated in the first iteration of the MVE in 2021–2022, Stage 2 describes the second iteration in 2023–2024, and Stage 3 the third iteration in 2025–2027.

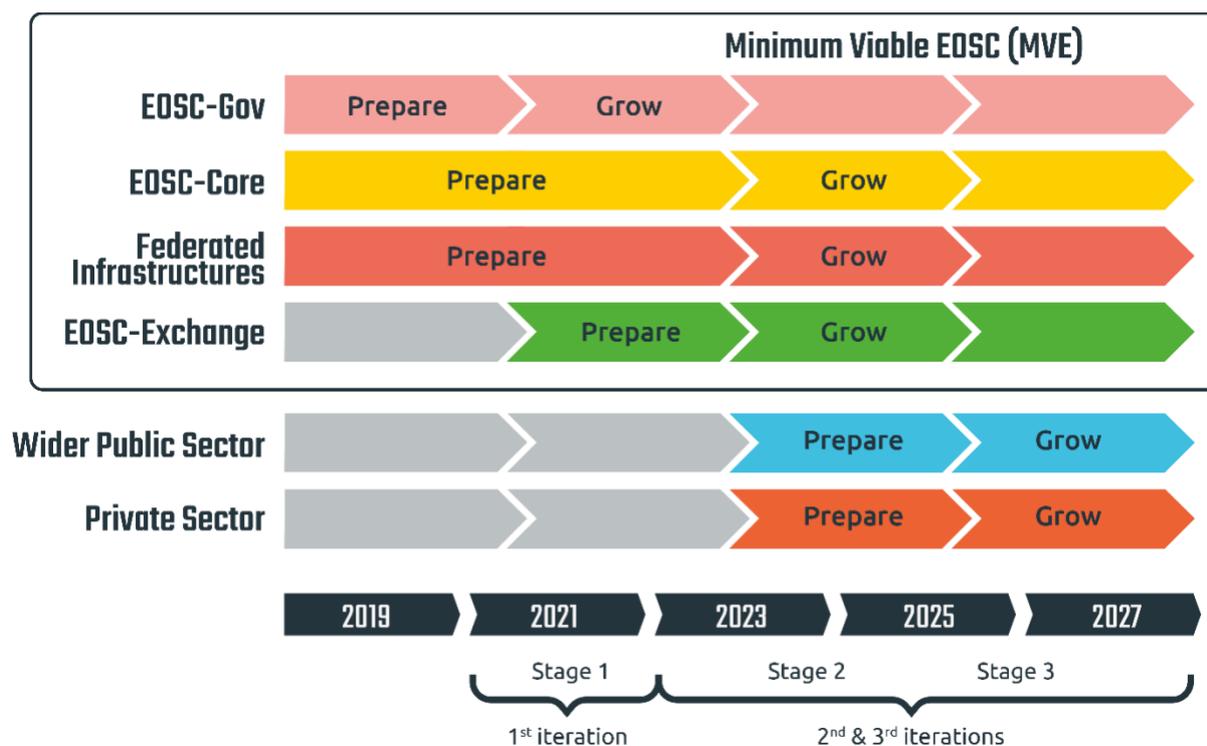


Figure 7. Scope and timing of the MVE

The EOSC governance has been well developed during the first stage with the establishment of the EOSC Association and the Tripartite Collaboration with the EC and Member States and Associated Countries through the EOSC Steering Board. The EOSC Association General Assembly now comprises 229 organisations, of which 155 are full Members and 74 are Observers. These members come from a wide range of countries and organisation types, including funders, universities, national facilities, commercial providers and not-for-profit bodies. Such representation enables the Association to act as the voice of the community. Moreover, a Secretariat of six members of staff has been established, including a Secretary General and Chief Technology Officer. The team will grow to 10 members of staff in 2022. As such, the Association is well placed to stimulate coordination across the funded EOSC implementation projects and represent the stakeholder community when steering development and implementation of EOSC. The Association’s thirteen Task Forces will play an instrumental role in setting the agenda and coordinating developments in specific areas of implementation. As EOSC continues to evolve, playing an active role in making it operational will require a further increase in Association staffing.

The development of EOSC-Core, EOSC-Exchange, and the federation of RIs is already underway. The EOSC Architecture description, laid out by previous Executive Board Working Groups, has been progressed in developing these components. Building blocks, such as the Interoperability Framework and the associated governance process, allows these procedures to be implemented. Building on the outcomes of previous projects, work has advanced to allow data catalogues to be harvested and research-graph technology is being used to connect the existing service and resource catalogues.

Activities in the roadmap for 2023-2024 focus on further growth to embed the EOSC-Core, EOSC-Exchange and governance structures, as well as to prepare to extend activity to the public and private sectors.

### **2.3 H2020 and Horizon Europe EOSC projects**

The implementation of EOSC is driven by a series of EC funded projects related to EOSC. Projects still active from the 2021 period are represented in Figure 8. Several of these were supported under the Horizon 2020 work programme, with those starting since 2022 being supported under Horizon Europe. The wide range and diversity of projects illustrates the challenge faced by the EOSC Association in terms of coordination. Attempting to build a coherent infrastructure via a series of projects is problematic, especially when their accountability and evaluation is based on individual work plans rather than the overarching EOSC SRIA. The proposed arrangement between the EOSC Association, the HE projects and the EC under Horizon Europe, to ensure collaboration and coordination with the EOSC Association, should alleviate this position. Realising EOSC will also require a willingness on the part of projects and the EC to be flexible and allow changes to proposed plans to avoid overlapping activities or divergent approaches, and instead jointly work towards an orchestrated result.

The INFRAEOSC calls are of particular importance since they support the major investment in building EOSC and fostering the development of FAIR and Open Science practices. Within this programme, a series of grouped projects have been supported to represent the science clusters, regional initiatives, and free-at-the-point-of-access services. These groupings and associated projects are listed below for reference. Collaboration agreements with these groups have been established, to foster alignment of joint activities.

#### **INFRAEOSC 04-2018 – Connecting ESFRI infrastructures through Cluster projects**

- Photon and Neutron Open Science Cloud (PANOSC)
- ENVIRONMENTAL RESEARCH INFRASTRUCTURES BUILDING FAIR SERVICES ACCESSIBLE FOR SOCIETY, INNOVATION AND RESEARCH (ENVRI-FAIR)
- Social Sciences & Humanities Open Cloud (SSHOC)
- European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures (ESCAPE)
- EOSC-Life

#### **INFRAEOSC 05b-2018 – Coordination of EOSC-relevant national initiatives across Europe**

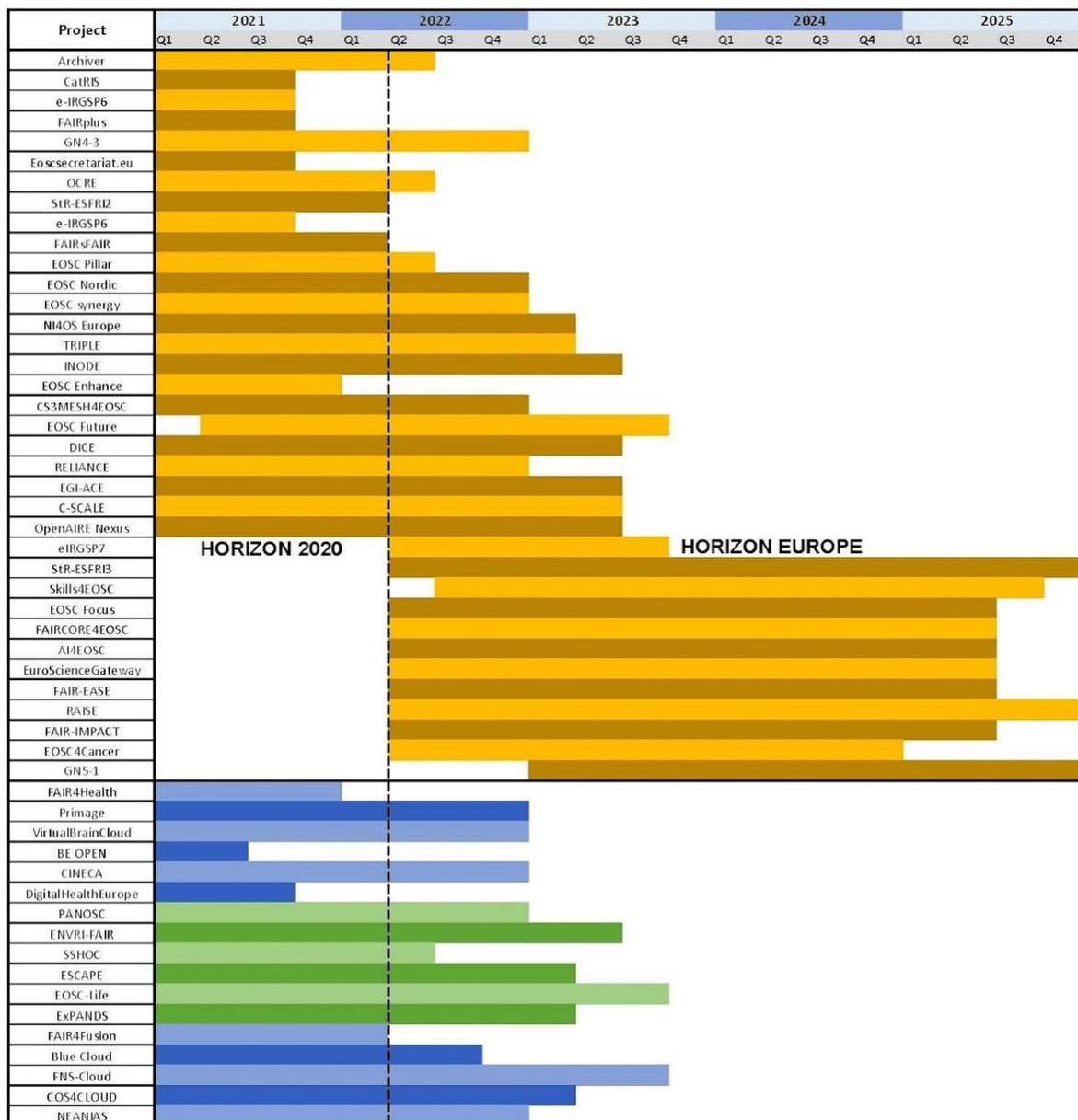
- Coordination and harmonisation of national initiatives, infrastructures and data services in Central and Western Europe (EOSC Pillar)
- EOSC-Nordic
- Expanding Capacities by building Capabilities (EOSC Synergy)
- National Initiatives for Open Science in Europe (NI4OS-EUROPE)
- EOSC Photon and Neutron Data Services (ExPaNDS)

#### **INFRAEOSC 07-2020 – Increasing the service offer of the EOSC Portal**

- Data infrastructure capacity for the European Open Science Cloud (DICE)
- RESEARCH LIFECYCLE MANAGEMENT FOR EARTH SCIENCE COMMUNITIES AND COPERNICUS USERS IN EOSC (RELIANCE)
- EGI Advanced Computing for EOSC (EGI-ACE)

- Copernicus - eoSC Analytics Engine (C-SCALE)
- OpenAIRE Nexus Scholarly Communication Services for EOSC users (OpenAIRE-Nexus)

It should be noted that many of these projects will end in the 2023-2024 period and new ones will begin. As such, a constant monitoring and engagement function is required to ensure that relevant activities across projects are aligning and continuing from one to the next. Funding infrastructure and service development via several standalone short-term projects is not the optimum mechanism to ensure robust and sustainable service delivery.



**Key:** Orange shades indicate coordination, best practice or technology projects. Blue shades indicate thematic science-based projects. Green shades indicate ESFRI cluster projects.

Figure 8. EOSC projects active since 2021

## 3. Roadmap 2023-2024

### 3.1 SRIA objectives

The Co-programmed European Partnership on EOSC, as agreed between the EC and the EOSC Association, is to be implemented in an open, transparent, efficient, and flexible way. The EOSC Association demonstrates this openness and transparency via the participatory approach in which we consult with stakeholders when developing key documents like the SRIA, and our inclusiveness when defining membership for structures like our Task Forces. Moreover the Memorandum of Understanding for the co-programmed partnership and other key documents are publicly available. The intended cooperative relationship aims to achieve jointly defined objectives based on a long-term common vision and a clear commitment from the partners throughout the duration of the Partnership. More on the EOSC Partnership can be found in Section 2.9.1 of the SRIA.

Slight edits have been made to the Specific and Operational Objectives since version 1.0 of the SRIA. These include removing all dates, since only some objectives included timeframes, and the delivery is tracked via the KPIs rather than the objectives. Also, periphery text has been removed to assist in readability. The edits made did not change the meaning or emphasis of any objective. The only significant change was to add Specific Objective (SO) 10. This relates to a new priority area of work, namely data quality, which was not reflected in the roadmap previously.

The **general objectives** (GOs) of the European Partnership, which are identical to the strategic objectives described in Section 3 of this SRIA, are defined as follows:

- GO1** Ensure that Open Science practises and skills are rewarded and taught, becoming the 'new normal';
- GO2** Enable the definition and adoption of standards, and the development of tools and services, to allow researchers to find, access, reuse and combine results;
- GO3** Establish a sustainable and federated infrastructure enabling open sharing of scientific results.

The **specific objectives** (SOs), which are reflected in the critical success factors identified in Section 7.4 of this SRIA, are the following:

- SO1** Increase in the number of relevant research outputs that are made available as open as possible by researchers performing publicly funded research;
- SO2** Professional data stewards are increasingly available in RPOs in Europe to support Open Science;
- SO3** Development and adoption of incentives for researchers to perform Open Science;
- SO4** Increasing amounts of research outputs from publicly funded research in Europe are FAIR by design;
- SO5** The EOSC Interoperability Framework supports an increasing range and quantity of FAIR digital objects including data, software and other research artefacts;
- SO6** Provide an increased number of services and resources to ensure that European research is discovered and reused within and across disciplines to extract new knowledge;

**SO7** EOSC is operationalised and provides a stable and valuable infrastructure supporting researchers to address societal challenges;

**SO8** Essential additional functionalities for end users from the public and private sectors are implemented in EOSC (these developments are complementary to those of other European data spaces);

**SO9** EOSC increasingly establishes ties with related initiatives from regions around the world and becomes a partner in global cooperation frameworks for Open Science;

**SO10** Common data quality indicators are agreed and implemented to ensure that research outputs within EOSC are ready for FAIR usage.

The **operational objectives** (OOs), which are reflected in the action areas described in Sections 5 and 6 of this SRIA, are the following:

**OO1** Deliver and operate all the necessary components of the MVE to openly share research data, publications, software, tools and services while attracting increasing numbers and categories of users, including those from public and private sectors.

**OO2** Develop and evolve monitoring systems to gather data and evidence on best Open Science practises accessible through EOSC. This could utilise the EOSC Observatory dashboard to monitor the evolving landscape of policies, infrastructures and open resources;

**OO3** Increasingly mainstream Open Science skills in European RPOs, including through the uptake and alignment of curricula and training frameworks related to the full lifecycle of data management;

**OO4** Co-develop domain-specific standards and adopt Open Science practises through engagement with research communities;

**OO5** Provide the technical components of a FAIR ecosystem for uptake and customisation by the communities (including open specifications, standards, schemas, application programming interfaces (APIs), metadata frameworks supporting FAIR digital objects and their automated processing);

**OO6** Provide the metrics and tools to measure the adoption of the FAIR principles for research outputs and provide frameworks to help in certifying that repository services enable FAIR;

**OO7** Co-develop a first generation of a robust pan-European network of infrastructures for software source code (including incentives for the effective documentation and sharing of research software);

**OO8** Co-design and adopt a rewards and recognition framework for FAIR and open science practises in research;

**OO9** Implement and evolve the EOSC Rules of Participation and onboarding process for EOSC providers and increase the number of service providers and services offered progressively;

**OO10** Deploy and operate an AAI framework to manage user identity and access;

**OO11** Implement the EOSC PID policy and architecture, including the development of a global PID resolver;

**OO12** Co-design a minimum metadata framework and provide a common search mechanism to EOSC resources;

**OO13** Continuously monitor and promote the increased uptake of core services and EOSC resources, access to EOSC-Exchange tools and services and ensure a feedback loop with the users;

**OO14** Define models for availability and costing of services across borders.

### 3.2 Levels of implementation

The activities for implementing all stages of this MAR are outlined below (see Sections 3.4 to 3.6). Those carrying out these activities can best be seen as operating at three levels of implementation:

- **European level (L1):** everything done at the European level either by or financed through the Horizon Europe programme or other sources, be it an effort by the EOSC Association or by a research infrastructure or service organisation operating at a European level. This could also be achieved through the joint effort of countries, etc. In other words, every contribution to EOSC being worked on at the European level (irrespective of who pays).
- **National level (L2):** the same as above but then at the country level, i.e. activities in one of the Member States or countries associated to the Horizon Europe programme (MS/AC), as long as the activities contribute positively to the development of the EOSC ecosystem as described in this SRIA.
- **Institutional level (L3):** again the same as above but then the activities at the level of the participating institutions (e.g., a university or other RPO), and again as long as the activities are aligned with the country strategy which should in turn be aligned with the European strategy as described in the SRIA.

### 3.3 Summary of priorities

The priorities for 2023-2024 are drawn from the evolution of the 2021-2022 roadmap, the forthcoming EC work programme, the EOSC Association workplan and responses from the public consultation. They focus on embedding as much as possible the foundational governance structures and EOSC technical infrastructure achieved in 2021-2022. The consultation suggested a number of priority areas of work which have been agreed by the EOSC Association Board and are explained in the paragraphs below, namely:

- Business / resourcing models
- EOSC-Core development and onboarding procedures
- Interoperability and data search
- Data quality
- Member State engagement
- Recognition and rewards

The key priority in the 2023-2024 period is to identify and test a number of funding models to ensure the EOSC-Core and EOSC-Exchange are viable over the long-term. Each component will be funded differently; the Core will be supported in a common manner by public funds, while the large variety of services in the Exchange will be resourced using different mechanisms. The EOSC-Exchange must accommodate the range of funding models in use by service providers who wish to federate into EOSC. Such funding models could include subscriptions, licence payments or other pay-per-use models, as well as services that are

made freely available. In addition, the Exchange should enable publicly-funded infrastructures such as data repositories and computational resources, to be used in a cross-border context with the possibility of recovering costs. This area of work is complex and it is critical to prioritise efforts here in order to clarify the value proposition for any service provider to federate into EOSC.

Associated with the business / funding models, emphasis is placed on solidifying the EOSC-Core and onboarding procedures. It must be clear how to federate into EOSC and to have efficient processes to enable RIs and other service providers to expose their data and services via EOSC. There are parallel developments here in the 2023-2024 period since the EOSC Future project is scheduled to end in September 2023 and the public tender for the EOSC procurement call will likely be awarded in early 2023. The EOSC Association Board strongly recommends that the tender specification requires that the awarded contractor builds upon the EOSC-Core components and standards that are widely used in the research community and that cooperation agreements are signed between the contracting parties, the EOSC Future project and the EOSC Association to ensure collaboration towards an agreed common goal.

The EOSC-Core addresses interoperability as one of its four key components (see Figure 6) and since this is the glue that enables the federation of data and services, much more emphasis and effort needs to be directed to these activities. The implementation projects must adopt existing open standards, open interfaces, open APIs, and promote the use of crosswalks as a way to ensure an interoperable and functional EOSC-Core. The EOSC-Core functionality also includes a common search for content (i.e. data, software and other resources) held within EOSC. This is at a fledgling state currently but must be improved to ensure value is demonstrated to research communities and that greater reuse of research outputs is enabled. The search functionality should adopt accepted standards (e.g. DDI-CDI, DCAT, schema.org) and support crosswalks to discipline-specific metadata standards that are in use by the RIs and the scientific communities. In order for this to succeed, community engagement on agreeing implementation approaches and the role for the EOSC Task Forces in steering what is happening within projects is paramount.

The EC emphasised data quality as a new area in the forthcoming work programme, in line with EOSC Association suggestions. While policy drivers push for research outputs to be Open and FAIR, content that is readily available and reusable is useless if it is of poor quality. Work is needed to define common data quality indicators for EOSC and to implement these.

Many questions were asked about the resourcing of EOSC. Much investment has been made by the EC into core infrastructure and services, but given that so much research infrastructure is funded nationally, greater Member State engagement is also required. Workable business models are a prerequisite however for Member States to co-invest and encourage or require national infrastructure and services to federate into EOSC.

Another area of priority is that of recognition and rewards. The way we evaluate researchers has been changing towards rewarding Open Science, but this needs to intensify. Incentivising open science practices needs to be made the norm. The EC, other funders, and research institutes will have to play a huge and important role in this.

In the following three sections, priorities are structured according to the objectives listed in the SRIA. For each objective a series of priority activities and expected outcomes is specified, as well as contextual information to explain plans for implementation or potential issues that need to be highlighted.

### **3.4 Objective 1 - Open Science as the new normal**

Objective 1 states:

**Ensure that Open Science practises and skills are rewarded and taught, becoming the ‘new normal’**

Below are the priority activities needed to enact this change in the 2023-2024 period at European, National and Institutional level.

#### *3.4.1 European level priorities*

- Include Open Science principles in EU data policy and legislation, as has been done by incorporating access to data in the Open Data Directive. The Open Data Directive should be transposed into Member State law, including the non-use of sui generis Database Protection Rights by public sector bodies. This is an important step towards broader legal interoperability of resources.
- Liaise with Horizon Europe projects and other key EOSC projects to coordinate activities jointly leading to a MVE.
- Manage and evolve the EOSC Observatory to record statistics on investments, policies, research outputs, open science skills, and infrastructure capacities being planned and delivered via Member States, Associated Countries, and EOSC Association members.
- Clarify the overall rules for onboarding into the Minimum Viable EOSC and produce simple manuals for service providers to follow. Also align the existing procedures for onboarding with the EOSC Rules of Participation.
- Provide support for users from different target groups to federate their digital environments with those of other national providers, horizontal and thematic RIs, and e-Infrastructures.
- Provide ample opportunities for the community to learn about, be consulted on, and engage in EOSC developments under the leadership and coordination of the EOSC Association.
- Widen and deepen the possible EOSC user base via collaboration with various relevant stakeholders.
- Focus on engaging research communities to increase their participation in EOSC.
- Continue to support the development of curricula for Open Science and FAIR skills and align these at a European level to ensure a base-level standard is reached across the European research community.
- Support the development of networks for data stewards, research software engineers, and Open Science communities at European level.
- Stimulate the definition of measures for researcher assessment to help create better rewards and recognition systems for quality Open Science. Although it has a broader remit than recognition and rewards for Open Science, the EC’s research assessment group is a useful vehicle for progressing this activity.

- Support the creation and dissemination of shared domain specific standards and evaluation frameworks and how they can be recognised, e.g., via open badges.
- Develop the EOSC visual identity and define the conditions for its reuse by stakeholders, based on the new EOSC logo.
- Systematically review and evaluate the impact of new metrics, checking they are incentivising open science practice and mitigating any unintended effects.

### *3.4.2 National level priorities*

- Demonstrate national commitment to EOSC by actively contributing to activities of the EOSC Steering Board, the Tripartite Collaboration, and the Council of National Open Science Coordination, including the sharing of best-practices.
- Discuss long-term funding commitments and the time after Horizon Europe (2028 onwards) on a national level.
- Harmonise national Open Science monitoring and KPIs to feed into the EOSC Observatory activities.
- Focus on engaging research communities to increase their participation in EOSC.
- Coordinate awareness raising and uptake of EOSC in collaboration with the designated EOSC Association Mandated Organisations or other EOSC national structures, national funders, and RIs to interlink the national Open Science ecosystems to EOSC.
- Develop national policies and funding streams to support the development of EOSC and incentivise national RIs to federate into EOSC.
- Support regional corporations across countries (e.g., the Nordics) learning from regional projects to strengthen national engagement.
- Incentivise and encourage smaller community and institutional infrastructures in their growth towards ensuring FAIR data and providing long term access.
- Establish the cost of data management, data stewardship and maintenance of research outputs (including software) as eligible within national funding schemes.
- Develop curricula for Open Science and FAIR data skills in line with aligned European curricula and encourage the inclusion of these as a core element in research programmes.
- Support training of data stewards and research software engineers, increasing their numbers in the European scientific context.
- Offer Open Science and FAIR skills training to policy makers, funders and other relevant stakeholders, like evaluators to promote alignment in evaluation criteria.
- Establish policy and coordinate implementation frameworks that measure and reward FAIR and Open Science practises.
- Recognise research data and software properly as outputs for research in national evaluation of science.

### *3.4.3 Institutional level priorities*

- Focus on engaging research communities to increase their participation in EOSC through, e.g., Rectors' conferences, EOSC reference points, and national and regional tripartite events.
- Incentivise and encourage institutional infrastructures to share best practises on how to channel a broad span of researcher needs into the further development of EOSC.

- Offer core Open Science and FAIR training to researchers at all levels and recognise these skills as important for professional development. Such training could be incorporated into doctoral training programmes together with research integrity.
- Embed EOSC information into institutional research programmes to support uptake by research communities.
- Foster awareness of local FAIR-enabling practices and the value of long-term preservation of data and metadata amongst research communities.
- Support professional development programmes to ensure research support staff have the required Open Science and FAIR skills.
- Engage in networks for data stewards, research software engineers, and Open Science communities to share models and harmonise best practises.
- Provide meaningful career pathways that allow research software engineers and data stewards to progress in academia from postdoc level and beyond.
- Support the development of an open science policy and associated implementation groups to create an open institutional culture.
- Adjust research review mechanisms and researcher incentives to ensure FAIR research outputs and Open Science are appropriately recognised and rewarded.

#### *3.4.4 Expected outcomes*

- A growing number of stakeholders actively engaged in EOSC, representative of the various groups and disciplines at European, national, and institutional level, that channels EOSC information to the community, supports alignment, and helps to ensure EOSC developments meet real-world use cases.
- FAIR, persistent and trusted research outputs and services from a growing range of research communities are available in EOSC.
- Validated statistics on national and institutional contributions to EOSC are available, supporting future strategy development.
- Researchers, including doctoral candidates, are equipped with the appropriate Open Science and FAIR skills and are supported by highly skilled professionals to transform the way they use research outputs, leading to better-quality and robust research.
- Researchers are supported by incentives for career advancement that include open science principles.
- Data Stewards, Data Scientists, Research Software Engineers and other associated professions are established as new job profiles in the science system.
- A higher proportion of research software engineers and other highly-sought staff remain in academia as their skillset is appropriately recognised and rewarded.
- Data stewards and researchers can network at a European and national level to access peers for support, consolidating good practices.
- After having established the new EOSC visual identity, the rules for using the EOSC brand are codified, communicated, and consistently applied.
- Open Science becomes a national priority and the public national and regional funding agencies use Open Science principles as a funding condition for research activities.

#### *3.4.5 Clarification text*

Coordination across EOSC implementation projects is paramount to ensure developments result in a coherent, operational platform. The EOSC Association will play a key role here. Hence, all new Horizon Europe projects should be encouraged to collaborate and engage with the EOSC Association and to contribute to this SRIA. As coordinator of the EOSC Focus project, the Association will oversee these activities by facilitating the liaisons and the alignment among the EOSC projects and the relevant stakeholders acting at European, national and institutional level by setting up an EOSC community platform and hosting regular coordination and alignment fora. These activities will provide the foundations for the planned liaison, necessary to effectively co-create EOSC, ensuring that the needs of research communities are appropriately addressed and to widen EOSC engagement. The EOSC Focus project will also continue the monitoring function. For this to work, the EOSC Association must be positioned as the “Data Controller” of the monitoring data.

The role of national structures is also key to incentivise RIs to federate into EOSC and support smaller communities on their journey towards FAIR and Open Science practices. While much work has already been done to align EU and national policies with Open Science principles, actions are needed to incentivise the adoption of these practices. This will include support for development of standards and a network of FAIR-enabling trusted repositories.

Following on from the definition of the Rules of Participation by the EOSC Executive Board RoP Working Group in 2021, a review should be undertaken to ensure the onboarding procedures implemented by the projects align with these Rules and are clear and fit for purpose. This activity should evaluate whether the expected number of national RIs and community services have been federated into EOSC to provide a solid platform from which to build. There should also be clear added value for service providers (e.g. integration into EOSC AAI, access to monitoring etc) and those offering data and code (e.g. increased discovery, ability to combine data from different domains) to federate into EOSC. Actions at European level and from Member States are needed to continue to incentivize and encourage participation amongst service providers and research communities to build the EOSC federation.

Since a growing number of courses are emerging for FAIR and Open Science, priorities in this phase of implementation focus on aligning curricula to ensure a base-level competency across the European research community. While the European level is responsible for coordinating activities and providing networks to enable data stewards, research software engineers, and Open Science communities to engage with peers, much activity focuses on the national and institutional level. It is within these bounds that core competencies can be included in research programmes to ensure a base level of skills. Moreover, national funders and research organisations operate the recognition and rewards programmes and set grant conditions which will be critical to incentivising researchers and support staff and to providing appropriate career paths to retain their expertise. Systematic reviews of the new metrics and incentives applied are also needed to ensure they have the desired effect. European level coordination is also needed, since fragmentation is still a serious problem when it comes to collecting and using new metrics.

Data stewards and research software engineers are largely new funded roles at the RPOs. Researchers may transition into these new roles, hence it is important to determine possible career paths for them. There is also a need to increase the number of professional support

staff, hence the need to specify those job roles and offer aligned curricula and support fora to ensure they are recognized and aligned across Europe.

### **3.5 Objective 2 - Adoption of standards and tools**

Objective 2 states:

**Enable the definition and adoption of standards, and the development of tools and services, to allow researchers to find, access, reuse, and combine results**

Below are the priority activities needed to enact this change in the 2023-2024 period at European, National and Institutional level.

#### *3.5.1 European level priorities*

- Embed the process by which community standards are endorsed for use within EOSC, developing multiple registries to address metadata standards and other semantic artefacts.
- Further enhance the EOSC Interoperability Framework to foster composability between services and resources, facilitating the creation of innovative solutions that enable research use cases and foster collaboration between diverse research communities.
- Develop and maintain open interfaces, alignments, crosswalks, and APIs that enable interoperability and foster adoption of EOSC. Research communities must be supported financially to undertake such work to facilitate interoperability.
- Establish clear interoperability guidelines towards all Data Spaces, foreseen in the European Data Strategy.
- Identify research communities and disciplines that lack standards and promote their development.
- Define, disseminate and implement common measures of quality to be applied to data, software, and other semantic artefacts, so the content within EOSC can be evaluated by users.
- Monitor technology developments and adopt new standards and software languages, where appropriate, to avoid that components of the EOSC-Core become outdated.
- Integrate EOSC services into EOSC-wide AAI systems. Core service integration should be mandatory whilst integration of EOSC-Exchange services should be strongly encouraged.
- Collaborate with national and thematic competence centres to support adoption of the EOSC-Exchange amongst service providers and researchers.
- Ensure the quality, long-term availability, and usability of the services in EOSC-Exchange by means of detailed specifications, standards adoption, assessment, certifications, and monitoring.
- Evaluate and, as necessary, develop trusted environments for managing, preserving, and sharing sensitive data, recognising work already in place at national level.
- Develop a global scalable PID-resolving infrastructure to support the broad ecosystem of PIDs used. The PID-resolving infrastructure will re-use existing mechanisms as much as possible.

- Support a network of FAIR-enabling trusted repositories for EOSC, including code repositories, that provide and adopt interoperable tools and aligned services. This network should act as a community platform for shared experience and practices that engages with and uplifts wider data services.
- Develop multilingual tools and search to facilitate uptake by researchers across Europe using their native languages.
- Promote the alignment of vocabularies not only in terms of disciplines but also in terms of cultural and linguistic coverage.

### *3.5.2 National level priorities*

- Encourage and incentivise effective documentation and sharing of all outputs relevant to research, including software and source code by defining dedicated funding mechanisms to support it.
- Provide national RIs and RPOs with sustainable funding streams related to EOSC, domain, and cross-domain initiatives.
- Create access to resources from public authorities within EOSC, engaging with national legal and ethical experts to promote appropriate sharing of sensitive data at a European level and alignment to facilitate cross-border access.
- Support and incentivize the development, maintenance and adoption of open standards and APIs to enable resource composability and to achieve interoperability across communities including in the private sector.
- Encourage research communities and national RIs to define their interoperability guidelines and, where relevant, develop crosswalks (i.e. mappings of relationships between fields of different metadata schemas) and alignments with others.
- Incentivise assessment and open certification of data repositories to ensure they enable FAIR outputs, are trustworthy and sustainable, and provide an interoperability layer across the European-wide network.
- Encourage and promote the different cultural dimensions and native languages in the development of tools and standards to maintain the diversity of European communities.
- Contribute to the alignment of national vocabularies in all the scientific disciplines with other languages and prepare their sustainabilities through regular updates.

### *3.5.3 Institutional level priorities*

- Raise awareness of infrastructure to support the creation, management, and sharing of research data and software, encouraging adoption.
- Support research communities to adopt both general and domain-specific standards to increase adoption of FAIR practices and reuse.
- Provide guidance and assistance to small repositories to enable sustainability and interoperability through standards compliance, assessment, and certification.
- Advocate for research communities to support and test the crosswalks and APIs being developed to support interoperability across and inside research disciplines and repositories.

- Support the adoption of common quality requirements for FAIR research outputs to be implemented at EOSC level, including through institutional, disciplinary, and other data repositories.
- Implement tools to plan, track, and assess scientific knowledge production (e.g., by connecting science knowledge graphs, Research Information Management Systems, Data Management Plans (DMPs), and metrics) based on open definitions, standards, and models.

#### *3.5.4 Expected outcomes*

- A wide range of disciplinary standards are endorsed and adopted in EOSC and the associated interoperability developments support a growing range of use cases on multi-disciplinary reuse of data and services.
- The adoption of metadata standards increases the proportion of FAIR data and the range of disciplinary resources that can be discovered via an EOSC search interface.
- Services offered within the EOSC-Exchange are reliable and of high quality, growing trust and adoption of EOSC.
- Infrastructure for sensitive data sharing is provided, via which a comprehensive collection of data from public authorities is made available for research reuse.
- A network of trusted data repositories is available and repository managers can identify clear benefits (e.g., efficiency and streamlined operations) of being part of EOSC.
- FAIR-enabling trustworthy repositories act as key nodes in the full data lifecycle network of institutional, national, and European practice.
- Transparency of (meta)data curation and preservation practice across FAIR-enabling data services including trustworthy repositories are mandated and incentivised.
- PIDs are used systematically and links between them are accessible via knowledge graphs to visualise and track research network activity.
- Common requirements for the quality of research software and data are agreed and these measures are used to define quality indicators and standards accepted by research communities.
- Infrastructures for research software and source code are promoted at national and institutional level, increasing the amount of software and code accessible via EOSC.
- Sustainability is secured by national (state and public) institutions and resources, not based mainly on participation in EU funding instruments (e.g., the HE programme).
- Shared best practices on the alignment of vocabularies and their sustainability are available and promoted.

#### *3.5.5 Clarification text*

One area of specific interest in the 2023-2024 period is extending the content available within EOSC to sensitive data from public authorities. Trusted environments for managing and sharing such data will be offered at the European level and will be supported by national level actions to facilitate uptake. A standard set of methods should be developed that can effectively enable data sharing and processing of sensitive content. These should be general enough to be applicable within the same country and in cross-border scenarios. Activities

should be cross-articulated with the Data Spaces foreseen in the European Data Strategy to strengthen integration with them.

FAIR-enabling repositories for data and code should be brought together in a network to support the development and adoption of standards, tools, and practices that promote interoperability and consistency of services. By sharing tools and practices across the network of repositories, less developed services can be supported and advanced. Such activities could facilitate automated metadata uplift and offer multilingual knowledge management tools and community curation mechanisms. Certification of repositories is also critical to ensure the services are trustworthy and can be relied upon to maintain the content available via EOSC. Work should also be undertaken to clarify how assessment frameworks for FAIR research outputs and FAIR-enabling trusted repositories work together. This could take into account disciplinary and generic initiatives, but also cultural and native language practices.

Mechanisms to assess the quality of data, code, and other semantic artefacts within the frame of EOSC are also needed to ensure that all stakeholders have a clear understanding of the adequacy of the information they are accessing for their scientific application. A data quality framework needs to be defined, which includes specifying quality indicators, exploring ways to disseminate quality information, providing clear standards to follow and considering potential certification schemes. Guidelines initiated in the EOSC Association Task Forces will be used to steer projects implementing these activities. Institutions will also play a key role to support the adoption of common quality requirements to be implemented within EOSC.

This second phase of work will also focus on growing the RIs and communities connected to EOSC and enhancing the interoperability across them. Collaboration with national and thematic competence centres is envisaged to increase the range of services within EOSC. Inputs from Member States and institutions are needed here to incentivise communities to engage and embed EOSC within a wider user group.

In terms of interoperability, the focus is on endorsing and, where necessary, extending existing community standards rather than inventing new ones for EOSC. By doing so, the different communities will have an abstraction layer which will make everything easier to interoperate with the EOSC-Core, and to share and compose resources with each other. The standards adopted should be broadly accepted by research communities, mature, well-maintained, and ideally in use globally. It is also critical to develop crosswalks and open APIs to allow different community standards to be integrated. A prioritisation is required for this integration work based on actual scientific use cases, for example the combination of social science and public health data to address the COVID-19 pandemic. It should also be noted that interoperability goes beyond semantics and addresses legal, organisational, and technical issues. For example, machine-actionable licences are needed to clarify rights management, addressing not only who has the permission to reuse content, but also who is responsible for storage, curation and preservation.

The long term sustainability of EOSC pivots around the role and mandate of the EOSC Association to oversee and steer project implementation. This mandate needs to be formally reflected in project work plans and consortium agreements, and procedures clearly defined for efficient execution. The mandate and role of the EOSC Association to endorse relevant project results, take responsibility for sustaining them where relevant, and govern the entire EOSC initiative should also be well-understood and recognised.

### 3.6 Objective 3 - Establish a sustainable and federated infrastructure

Objective 3 states:

#### Establish a sustainable and federated infrastructure enabling open sharing of scientific results

Below are the priority activities needed to enact this change in the 2023-2024 period at European, National and Institutional level.

##### 3.6.1 European level priorities

- Evolve the EOSC governance to effectively coordinate activities and sustain and grow the MVE.
- Widen the resources accessible for EOSC to cover the full spectrum of research outputs (e.g., models, software, and workflows) and reliable common services (e.g. AAI, accounting, and monitoring).
- Enable researchers across all disciplines and organisations (small and large research infrastructures and less developed organisations) easy access to the federated infrastructure.
- Evolve and further establish the architecture blueprint for the EOSC-Core to enhance the model of the EOSC federated architecture, and as a toolkit to federate research data and enable thematic data spaces in the EOSC-Exchange.
- Ensure the EOSC-Core has well-defined technical, governance, and sustainability plans, supported by assessments of critical elements that need high availability and reliability.
- Continue to evolve an EOSC AAI Federation to enable more communities to connect their AARC Blueprint compatible AAIs for seamless access to resources and services.
- Enhance the federated search functionality for EOSC to improve discovery, tailored recommendations, and reuse of content.
- Ensure a continuous security and data protection framework along GDPR and other relevant regulatory frameworks, addressing licensing and IPR.
- Support a study on sustainability models for long-term preservation services, which covers the costs of data storage, curation, and ensuring long-term usability.
- Provide federated long-term access and preservation infrastructure for EOSC, which can be utilised by repositories and RIs. This should include provision of services that support reproducibility of preserved results (e.g., workflows and data processing container images) as well as the data itself.
- Identify resourcing models that enable cross-border access and use of existing national infrastructure capacity.
- Trial resourcing models identified in Phase 1 to provide the availability of resources in the MVE, including use and contribution by the wider public and private sectors.
- Continue to liaise internationally to develop a global cooperation framework for Open Science infrastructures also using EC partnerships with other regions such as Latin America and Africa.

### *3.6.2 National level priorities*

- Align research and e-Infrastructure strategies at national level and also between the national and EU levels.
- Introduce measures to incentivise RIs and RPOs to adopt appropriate community standards to be federated into EOSC.
- Ensure national investments are directed to national RIs that adopt appropriate community and interoperability standards and can be federated into EOSC.
- Gather information on the financial models implemented by national and European RIs, national research and education networks (NRENs), and communities to understand any constraints on pooling resources and making services available at a European level to develop cost-recovery models that incentivize the participation of institutions.
- Continue to harmonise requirements for DMPs (which should address software and other outputs) and encourage the use of machine-actionable DMPs, so that content from DMPs can be more easily aggregated and exchanged with other systems.

### *3.6.3 Institutional level priorities*

- Encourage the adoption and use of the federated infrastructure to be provided by EOSC, by sharing examples of how research use cases have been addressed and signposting relevant resources and services when engaging with researchers.
- Encourage the use of domain specific and existing RIs.
- Define and implement procedures to select data, software, and other research outputs of long-term value to be preserved via EOSC.
- Adopt machine-actionable DMP tools and openly share institutional DMP data, where possible, in pursuit of establishing best practices and increasing FAIR alignment. DMPs should be interlinked to trusted repositories and other institutional internal systems for managing projects.

### *3.6.4 Expected outcomes*

- A legal structure for EOSC with a proper governance such that the MVE can be sustained and further developed, by implementing adequate Rules of Participation.
- Sustainability of EOSC is secured by co-investment at European and national levels with cost-recovery models that incentivize participation of institutions and cross-border activity.
- An increasingly broad range of content is available via EOSC and usage statistics demonstrate increased discovery and reuse,
- The operation of EOSC-Core services (such as AAI) and federated infrastructure model enable a wide range of communities to connect to EOSC.
- Consolidation and further enhancement of the EOSC-Core, strengthening the federation aspects and integration.
- Long-term preservation infrastructure is offered within EOSC, enabling selected research outputs to be retained and actively preserved for at least 10 years.
- Researchers are supported to select data, code, and other outputs of long-term value.

- Resourcing models to sustain the MVE are trialled and advanced, and a core set of storage and compute resources are offered cross-border.
- Data from DMPs is reused to share good practices, increasing adoption of FAIR, and making the management of research projects more effective.

### *3.6.5 Clarification text*

Within the area of long-term preservation, it is recommended to undertake further studies that can elaborate on sustainability models that go beyond individual organisations. This should investigate the costs and facilities needed, review which file and data formats are appropriate to support, consider different models of implementing preservation services and assess the relative cost/benefit ratio of each, and make recommendations on the minimum retention and preservation periods for different types of data and research software. Decision-making mechanisms to select which data to preserve are also critical and should be elaborated and implemented in conjunction with data services, research communities, and RPOs. Recommendations on long-term preservation services should be implemented by providing effective and cost efficient long-term storage infrastructure at European level to which different repository types can connect.

A critical challenge in developing EOSC is to move beyond project-based partnerships and collaboration models to more sustainable long-term operations, such that resource providers can offer resources to any researcher in Europe and are assured that the resources consumed by researchers from outside their targeted user community are consumed in a financially sustainable way. The EOSC Association Task Force on Financial Sustainability, together with the EOSC Focus project, will explore resourcing models that enable cross-border access and use of services. There is not expected to be a single business model that can support all of EOSC, and the financing for the EOSC-Core will differ from that of the services made available via EOSC-Exchange. The financing models of EOSC-Core should ensure its wide availability and should not be unduly influenced by any single supplier.

The governance of EOSC must work towards financial models that support federated, cross-border access to existing national capacity to avoid fragmentation and maximise the investments made across Europe. It is likely that the financing model of EOSC will evolve over time and will also be linked to the evolution of the EOSC Association. In the long term, basic infrastructure such as that required to operate the EOSC-Core should be financed by public money, whilst the EOSC-Exchange will implement a broader range of financial models which are yet to be defined. However, in the initial setup phase, stimulation for both the EOSC-Core and EOSC-Exchange should come from European and national funding.

There is an overarching focus under this objective on further developing the federated infrastructure and interoperability as an enabler of Open Science practices. Work should continue to enhance and embed the EOSC-Core, enabling a greater range of services and research communities to connect. The common digital search functionality should continue to be enhanced to improve search options and filters. Actions on national and institutional levels will have to support this by ensuring Research Infrastructures and communities adopt the appropriate standards to be federated into EOSC and make resources discoverable and reusable. Where communities lack standards to contribute to EOSC, initiatives should be supported to promote their development. The role of bodies such as the Research Data



Alliances are important here to link up with international stakeholders and form the European component of the global cooperation frameworks for Open Science.

## 4. Conclusion

The MAR for the 2023-2024 period takes into account the activities planned for the first stage and those forthcoming under the Horizon Europe INFRAEOSC Work Programme. It provides a framework of priorities with derived ambitious activities and clear indicators for the second stage of activities. The MAR content has been shaped by the EOSC Association Task Forces and a public consultation held in March 2022. As such it represents the view of the stakeholder community at large.

The implementation of EOSC is driven by a series of EC funded projects related to EOSC. The wide range and diversity of projects illustrates the challenge faced by the EOSC Association in terms of coordination. Expectations placed on new Horizon Europe projects by the EC to ensure collaboration and coordination with the EOSC Association, will help to alleviate this position. Realising EOSC will also require a willingness on the part of projects and the EC to be flexible and allow changes to proposed plans to avoid overlapping activities or divergent approaches and favour joint work towards an orchestrated result.

Next to this, the technical and financial challenges that have to be overcome in trying to reach the objectives are immense. Significant challenges lie in combining local, national, and regional initiatives towards a true Open Science Commons, with global convergence on standards that are truly able to improve research. The federated approach requires alignment and convergence on standards and practices, which is not only technically challenging but necessitates lengthy consensus-building processes to ensure acceptance and adoption by research communities.

As noted in Section 3.3, defining, testing, and implementing viable business and resourcing models to ensure that the EOSC-Core and EOSC-Exchange can operate effectively is the key priority for 2023-2024. Without this, the EOSC Portal (or whatever front end continues to be provided) will merely serve as a shop-front with little inside, since the incentives and requirements for recovering the cost of service provision will not be met. Building EOSC is less of a technical challenge than addressing the social and organisational issues linked to its finance and governance.

Fortunately, the EOSC governance has progressed during the first stage and is rapidly becoming accepted as the voice of the community with a General Assembly that now comprises 229 organisations. The Association is also well positioned to help coordinate across implementation projects and perform the necessary liaison with Member States to stimulate investments in EOSC. Ensuring political support will be key in the 2023-2024 period to ensure national open science initiatives and RIs are incentivised to federate into EOSC so the platform gains critical mass and can become a valuable resource for the research communities of the future.